Received: 26 January 2023 | Accepted: 7 September 2023

DOI: 10.1111/ene.16077

european journal of neurology
the official journal of the european academy of neurology

ORIGINAL ARTICLE

Association of magnetic resonance imaging phenotypes and serum biomarker levels with treatment response and long-term disease outcomes in multiple sclerosis patients

Luciana Midaglia[1] | Alex Rovira[2] | Berta Miró[3] | Jordi Río[1] | Nicolás Fissolo[1] | Joaquín Castilló[1] | Alex Sánchez[3] | Xavier Montalban[1] | Manuel Comabella[1]

[1]Servei de Neurologia-Neuroimmunologia, Centre d'Esclerosi Múltiple de Catalunya (Cemcat), Institut de Recerca Vall d'Hebron (VHIR), Hospital Universitari Vall d'Hebron, Universitat Autònoma de Barcelona, Barcelona, Spain

[2]Section of Neuroradiology, Department of Radiology, Hospital Universitari Vall d'Hebron, Universitat Autònoma de Barcelona, Barcelona, Spain

[3]Unitat d'Estadística i Bioinformàtica, Institut de Recerca Vall d'Hebron (VHIR), Barcelona, Spain

Correspondence
Manuel Comabella and Luciana Midaglia, Unitat de Neuroimmunologia Clínica, Cemcat, Hospital Universitari Vall d'Hebron, Pg. Vall d'Hebron, 119-129 08035 Barcelona, Spain.
Email: manuel.comabella@vhir.org and lmidaglia@cem-cat.org

Abstract

Background and purpose: The aim was to evaluate whether magnetic resonance imaging (MRI) phenotypes defined by inflammation and neurodegeneration markers correlate with serum levels of neurofilament light chain (NfL) and glial fibrillary acidic protein (GFAP) in relapsing–remitting multiple sclerosis (RRMS) patients; and to explore the role of radiological phenotypes and biomarker levels on treatment response and long-term prognostic outcomes.

Methods: Magnetic resonance imaging scans from 80 RRMS patients were classified at baseline of interferon-beta (IFNβ) treatment into radiological phenotypes defined by high and low inflammation and high and low neurodegeneration, based on the number of contrast-enhancing lesions, brain parenchymal fraction and the relative volume of non-enhancing black holes on T1-weighted images. Serum levels of NfL and GFAP were measured at baseline with single molecule array (Simoa) assays. MRI phenotypes and serum biomarker levels were investigated for their association with IFNβ response, and times to second-line therapies, secondary-progressive MS (SPMS) conversion and Expanded Disability Status Scale (EDSS) 6.0.

Results: Mean (SD) follow-up was 17 (2.9) years. Serum NfL levels and GFAP were higher in the high inflammation ($p=0.04$) and high neurodegeneration phenotypes ($p=0.03$), respectively. The high inflammation phenotype was associated with poor response to IFNβ treatment ($p=0.04$) and with shorter time to second-line therapies ($p=0.04$). In contrast, the high neurodegeneration phenotype was associated with shorter time to SPMS ($p=0.006$) and a trend towards shorter time to EDSS 6.0 ($p=0.09$). High serum NfL levels were associated with poor response to IFNβ treatment ($p=0.004$).

Conclusions: Magnetic resonance imaging phenotypes defined by inflammation and neurodegeneration correlate with serum biomarker levels, and both have prognostic implications in treatment response and long-term disease outcomes.

KEYWORDS
long-term prognosis, magnetic resonance imaging, multiple sclerosis, neurofilament light chain, treatment response

INTRODUCTION

Multiple sclerosis (MS) is a chronic immune-mediated disease of the central nervous system whose pathogenesis is mainly characterized by two major processes, inflammation and neurodegeneration [1, 2]. Clinically, the disease has a high degree of interindividual and intraindividual variability, and its unpredictable course makes disease management difficult, especially when it comes to making treatment decisions [3, 4].

Classically, MS monitoring relies on the assessment of relapses and disability measured by the Expanded Disability Status Scale (EDSS) [5]. In addition, periodic follow-up magnetic resonance imaging (MRI) examinations are performed to assess focal inflammatory activity, primarily defined by the presence of gadolinium (Gd) enhancing T1 or new/enlarging T2 lesions [6]. However, these markers of disease activity often fail to predict individual relapse rate, disability progression and therapy response [7]. Furthermore, MRI markers of neurodegeneration and prediction of conversion to secondary progressive MS (SPMS) are difficult to implement in clinical practice [8]. In this context, there is an urgent need to identify and validate biomarkers that could be used as surrogate measures for clinical end-points in a more individualized manner [9, 10].

In a previous study conducted by our group [11], a cohort of 108 relapsing–remitting MS (RRMS) patients was classified at baseline for interferon-beta (IFNβ) treatment into four radiological phenotypes defined by various degrees of inflammation and neurodegeneration in order to identify specific blood transcriptomic patterns associated with MRI phenotypes characterized by high and low neurodegeneration. Down-regulation of B-cell-specific genes in peripheral blood mononuclear cells and higher activation status in B cells were found from patients with high neurodegeneration phenotypes [11]. In the present study the first aim was to investigate whether inflammation and neurodegeneration MRI phenotypes were associated with serum levels of neurofilament light chain (NfL) and glial fibrillary acidic protein (GFAP), and secondly to evaluate the prognostic role of MRI phenotypes and serum biomarker levels on treatment response and long-term disease outcomes.

METHODS

Patients

From the initial cohort of 108 patients with RRMS who participated in our previous study [11], 80 patients were selected based on availability of serum samples in proximity to the baseline MRI scans. This cohort corresponded to RRMS patients who started immunomodulatory treatment with IFNβ as their first disease-modifying therapy (DMT). During follow-up, the presence of relapses and EDSS scores were recorded at regular in-clinic visits every 6 months. Brain MRI scans were performed annually during the first 2 years on IFNβ treatment, and then at the discretion of the clinician, to assess the presence of new or enlarging T2 lesions and Gd-enhancing lesions. In those patients who switched treatment during follow-up, MRI scans were also performed annually for the first 2 years on the new treatment and thereafter according to clinical disease evolution or type of treatment used. The study was approved by the Clinical Research Ethics Committee at the Vall d'Hebron University Hospital and all patients signed a written informed consent.

Baseline MRI phenotypes

Baseline brain MRIs were acquired on a 1.5 T superconductive magnet using a standardized protocol (2D fast spin-echo dual echo T2-weighted, and pre- and post-contrast [0.1 mmol/kg, 5 min delay] 2D spin-echo T1-weighted sequences) as previously described [11]. In all patients, two experienced neuroradiologists visually assessed the presence and number of Gd-enhancing lesions on post-contrast T1-weighted scans. Brain parenchymal fraction (BPF), a normalized brain volume measure, commonly used as a surrogate of whole-brain atrophy, was calculated on the pre-contrast T1-weighted scans using a fully automated segmentation technique. For calculating the non-enhanced T1 black hole volume an in-house automatic segmentation algorithm was used that measured the T1 lesion load from the initial T2 lesion segmentation that was used as lesion mask. T2 lesion segmentation was performed using a semiautomatic local thresholding contour technique (Dispimage, DL Plummer, University College, London, UK) or, if the lesion could not be outlined satisfactorily with this approach, by manual outlining. Relative T1 black hole volume or black hole fraction (BHf) is expressed as the ratio of T1 lesion volume to the T2 lesion volume.

Magnetic resonance imaging scans were first classified into low inflammation and high inflammation phenotypes according to the presence or absence of contrast-enhancing lesions. MRI phenotypes with high neurodegenerative component were defined as follows: (i) the presence of BPF values <0.83 or (ii) BPF values ≥0.83 and the presence of BHf values ≥10%. MRI phenotypes with low neurodegenerative component were defined by the presence of BPF values ≥0.83 and BHf values <10% [12].

Considering the two major pathological processes taking place in the central nervous system of MS patients, for the present study the four initial radiological phenotypes (low inflammation and low neurodegeneration; low inflammation and high neurodegeneration; high inflammation and low neurodegeneration; high inflammation and high neurodegeneration) were regrouped into two major phenotypes: inflammation and neurodegeneration. Inflammation phenotypes merged the two phenotypes with high inflammation and the two phenotypes with low inflammation. Similarly, neurodegeneration phenotypes merged the two phenotypes with high neurodegeneration and the two phenotypes with low neurodegeneration.

Quantification of serum biomarker levels

Blood was collected in proximity to the baseline MRI scans and IFNβ onset (Table S1). Briefly, peripheral blood was drawn by standard venipuncture and allowed to clot spontaneously for 30 min. Serum was obtained by centrifugation and stored frozen at −80°C until used. Levels of NfL and GFAP in serum were determined on the fully automated ultrasensitive Simoa HD-1 Analyzer (Quanterix), using the human NfL and GFAP assays purchased from Quanterix. Samples were run in duplicate diluted at a 1:4 ratio, and appropriate standards and internal controls were included in accordance with the manufacturer's instructions. The mean intra-assay coefficient of variation for duplicate determinations for concentration was 5% for NfL and 3% for GFAP. The inter-assay coefficient of variation was 8.7% for NfL and 6.9% for GFAP.

Disease outcomes

Radiological phenotypes and biomarker levels at baseline were investigated for their potential associations with the following disease outcomes.

(i) Response to IFNβ. Therapeutic response to IFNβ was evaluated by the time to evidence of disease activity (EDA). Disease activity was defined by the occurrence of at least one of the following situations: relapses; new or active lesions on brain MRI scans; and sustained increase in EDSS (1 point when the baseline EDSS was less than or equal to 5.5, and 0.5 points for baseline EDSS higher than 5.5). For the analysis, the time elapsed from IFNβ onset to the first manifestation of disease activity was considered taking into account the first 5 years on IFNβ treatment. Response to IFNβ was also evaluated following classification of MS patients according to two extremes of therapeutic outcome: patients without disease activity during the entire period of use of IFNβ treatment (IFNβ responders) and patients with a lack of response to a third DMT (IFNβ non-responders).

(ii) Time to second-line therapies and proportion of patients with second-line treatment at the time of the last follow-up visit. Time to second-line treatment was calculated as the time between the baseline and onset with second-line therapies. For the study, fingolimod, natalizumab, cladribine, alemtuzumab, ocrelizumab and rituximab were considered as second-line therapies. Anti-CD20 therapies used during follow-up to treat patients with SPMS were excluded from the analysis.

(iii) Time to SPMS and proportion of patients with SPMS at the time of the last follow-up visit. Time to SPMS was calculated as the time between the baseline and the time to develop a progressive phase of the disease defined by the presence of sustained progression of EDSS in the absence of relapses.

(iv) Time to EDSS 6.0 and proportion of patients with EDSS 6.0 at the time of the last follow-up visit. Time to EDSS 6.0 was calculated as the time between the baseline and the time to reach an EDSS of 6.0, by which the patient needs walking assistance. Time to EDSS 6.0 was confirmed at 6 months.

Statistical analysis

Associations between radiological phenotypes and biomarker levels were analyzed using Student's *t* tests. To assess the prognostic role of radiological phenotypes and biomarker levels for the different defined outcomes, event rates were calculated in person-years by dividing the number of observed phenotypes during the study period by the sum of all individual follow-up times. Survival estimates for the biomarkers were analyzed using Cox proportional hazards models; the Kaplan–Meier survival analysis was used to compare survival curves between the two radiological phenotypes. To assess the proportion of patients with second-line treatment and SPMS at the time of the last follow-up visit, chi-squared and Fisher's exact tests were performed. For the comparative analysis of the two extremes of therapeutic outcome, a Mann–Whitney *U* test or a Student's *t* test was carried out as appropriate for quantitative variables, and a chi-squared test for categorical variables. For all analyses, biomarkers were adjusted by age. Analyses were conducted using R Version 4.2.0. *p* values below 0.05 were considered significant.

RESULTS

Demographic and clinical characteristics of RRMS patients at baseline

Table 1 summarizes demographic and clinical characteristics of the whole cohort of RRMS patients at baseline of IFNβ onset and after segregation into radiological phenotypes. Mean age (SD) of the whole cohort was 34.1 (8.4) years and the female/male ratio was 3.0. Mean follow-up time was 16.7 (2.9) years. Inflammation phenotypes included 30 (37.5%) patients with high inflammation and 50 (62.5%) with low inflammation. Neurodegeneration phenotypes included 48 (60.0%) patients with high neurodegeneration and 32 (40.0%) with low neurodegeneration. Baseline variables such as sex, EDSS and number of relapses in the previous year were comparable amongst patients with high and low inflammation, and amongst patients with high and low neurodegeneration. Patients with high neurodegeneration were older ($p = 0.003$) and had longer disease duration ($p = 0.035$) at the time of IFNβ treatment onset. Mean follow-up time for patients with high and low inflammation was 16.2 (3.3) and 16.9 (2.7) years respectively, and for patients with high and low neurodegeneration 16.5 (3.5) and 16.9 (1.6) years, respectively (Table 1).

At baseline, serum NfL levels were not associated with disease duration, EDSS, number of relapses or follow-up time (Table S1). GFAP levels correlated with EDSS at baseline ($p = 0.02$) but not with the other baseline or follow-up clinical variables (Table S2).

Serum NfL levels are associated with the high inflammation phenotype and GFAP levels with the high neurodegeneration phenotype

First, the association between MRI phenotypes and serum biomarker levels at baseline was investigated. As shown in Figure 1a, serum NfL levels were significantly higher in the high inflammation phenotype compared to the low inflammation phenotype ($p = 0.04$), whereas no significant differences were observed between the high and low neurodegeneration phenotypes. In contrast, serum GFAP were significantly increased in the high neurodegeneration phenotype compared to the low neurodegeneration phenotype ($p = 0.03$), but levels were comparable between the high and low inflammation phenotypes (Figure 1b).

TABLE 1 Demographic and clinical characteristics of RRMS patients according to the MRI phenotypes.

		Inflammation phenotypes			Neurodegeneration		
Characteristics	**Whole cohort**	**High**	**Low**	***p* value**	**High**	**Low**	***p* value**
N (%)	80	30 (37.5)	50 (62.5)	–	48 (60.0)	32 (40.0)	–
Age (years)	34.1 (8.4)	32.5 (6.8)	35.1 (9.1)	0.194	36.4 (8.4)	30.8 (7.4)	**0.003**
Female/male (% women)	60/20 (75.0)	22/8 (73.3)	38/12 (76.0)	0.790	38/10 (79.2)	22/10 (68.8)	0.292
Duration of disease (years)	4.8 (5.2)	4.3 (4.7)	5.0 (5.5)	0.518	5.7 (5.6)	3.3 (4.1)	**0.035**
EDSS[a]	2.0 (1.5–2.5)	1.5 (1.5–2.0)	2.0 (1.5–3.0)	0.081	2.0 (1.5–2.6)	1.8 (1.0–2.1)	0.198
Number or relapses[b]	1.5 (0.7)	1.6 (0.9)	1.4 (0.6)	0.164	1.4 (0.7)	1.4 (0.8)	0.855
Follow-up time (years)	16.7 (2.9)	16.2 (3.3)	16.9 (2.7)	0.269	16.5 (3.5)	16.9 (1.6)	0.501

Notes: Data are expressed as mean (standard deviation) unless otherwise stated. Age and EDSS correspond to the baseline of IFNβ onset. Disease duration was calculated as the difference between disease onset and IFNβ treatment onset. Follow-up time was calculated as the difference between IFNβ onset and the time of the last visit. Significant *p* values are shown in bold.

Abbreviations: EDSS, Expanded Disability Status Scale; IFNβ, interferon-beta; MRI, magnetic resonance imaging; RRMS, relapsing–remitting multiple sclerosis.

[a]Data are expressed as median (interquartile range).

[b]Refers to the number of relapses in the previous year before IFNβ onset.

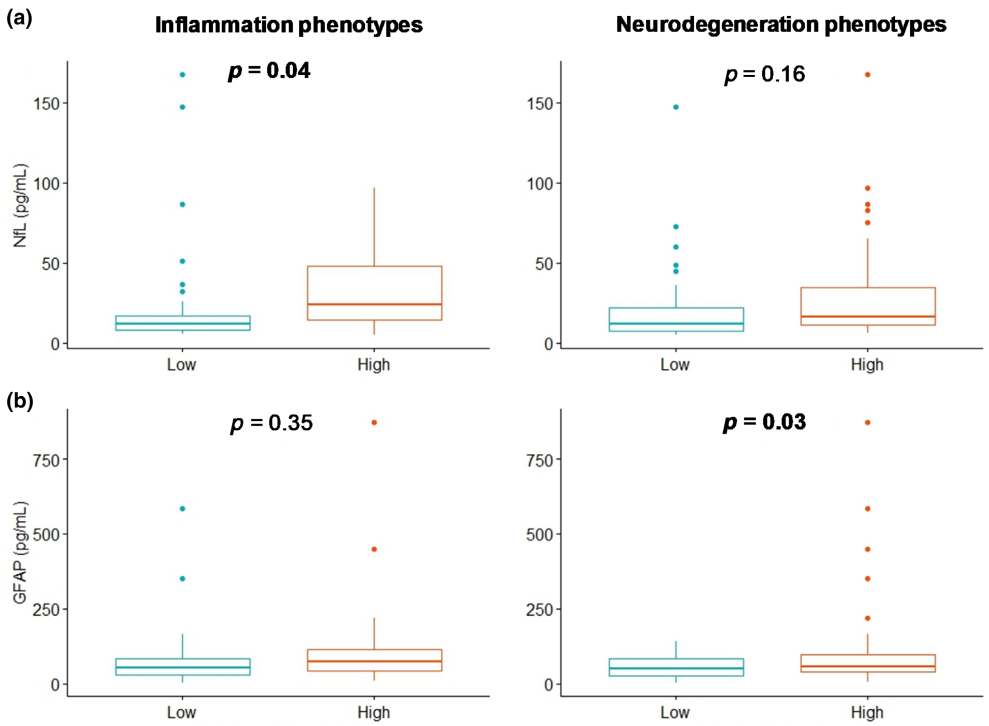

FIGURE 1 Association between radiological phenotypes and serum biomarker levels. Box plots showing the distribution of serum NfL levels (a) and GFAP levels (b) in MS patients with high and low inflammation, and with high and low neurodegeneration.

High inflammation phenotype and high serum NfL levels are associated with the response to IFNβ treatment

The association between radiological phenotypes and biomarker levels with the response outcome was next evaluated. As shown in Figure 2a, a shorter time to EDA was observed in patients belonging to the high inflammation phenotype compared to the low inflammation subgroup, with a median time to EDA of 1 year for the high inflammation phenotype and of 2.5 years for the low inflammation phenotype ($p = 0.04$). In contrast, time to EDA was similar between patients belonging to the high and low neurodegeneration phenotypes (Figure 2b).

High serum NfL levels at baseline were associated with an increased risk for EDA during IFNβ treatment (hazard ratio 1.014, 95% confidence interval 1.004–1.023; $p = 0.004$), whereas no association was observed between high serum GFAP levels and risk for EDA (Table 2).

Treatment response was also evaluated according to extremes of therapeutic outcome. Treatment responders ($N = 17$) were receiving IFNβ for a mean time of 10.2 (6.1) years, whereas non-responders to a third DMT ($N = 12$) were treated with IFNβ for a mean time of 4.5 (3.7) years ($p = 0.008$). On comparing the two extremes of therapeutic outcome groups, a trend towards significantly higher serum NfL levels was observed in the non-responder group ($p = 0.06$), whereas serum GFAP levels were comparable between the two groups of patients (Figure 3). The proportion of patients belonging to the high and low inflammation and neurodegenerative radiological phenotypes was also similar between IFNβ responders and non-responders to a third DMT ($p = 0.43$ and $p = 0.44$ for the inflammation and neurodegeneration phenotypes respectively) (data not shown).

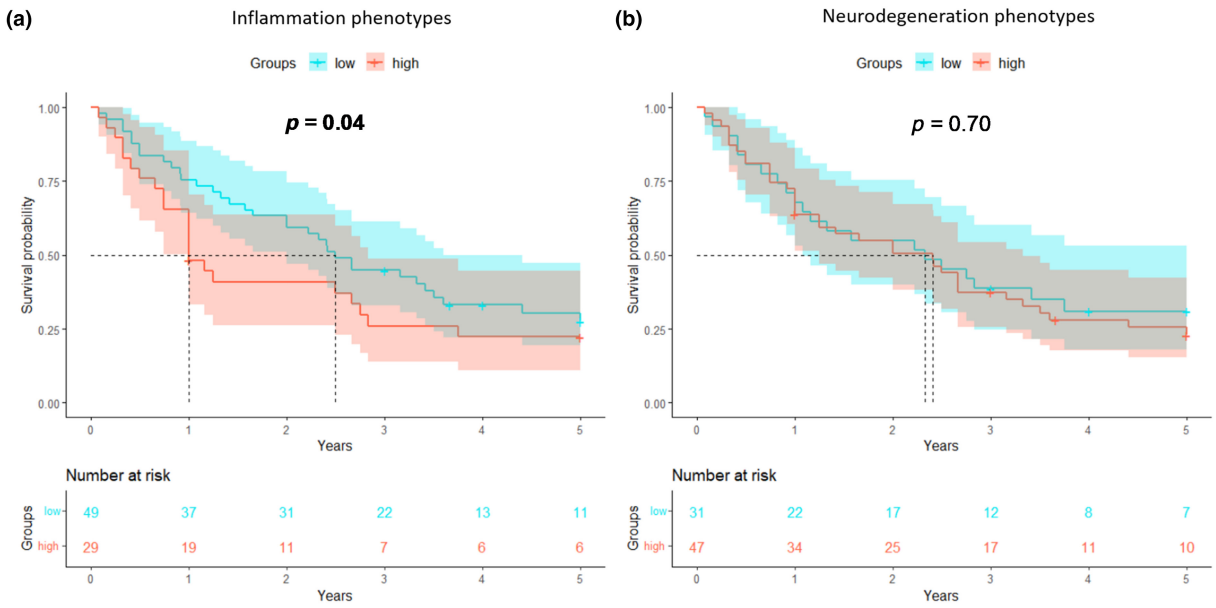

FIGURE 2 Time to evidence of disease activity at year 5 of IFNβ treatment. Kaplan–Meier curves showing the survival of patients with RRMS for the event evidence of disease activity in the first 5 years on IFN treatment for high and low inflammation phenotypes (a) and high and low neurodegeneration phenotypes (b). The blue and red lines correspond to survival probability for the low and high inflammation phenotypes, respectively. Shaded areas correspond to the 95% confidence interval for each curve, and overlap between confidence intervals is represented in gray. Discontinued lines indicate median times to the event for each group.

TABLE 2 Association between serum NfL and GFAP levels at baseline and disease outcomes.

	Serum NfL levels		Serum GFAP levels	
Outcomes	**HR (95% CI)**	***p* value**	**HR (95% CI)**	***p* value**
Time to EDA	1.014 (1.004–1.023)	**0.004**	1.001 (0.998–1.003)	0.41
Time to second-line therapy	1.006 (0.995–1.017)	0.26	0.998 (0.994–1.003)	0.59
Time to SPMS	1.000 (0.988–1.023)	0.95	1.001 (0.998–1.003)	0.67
Time to EDSS 6	0.986 (0.960–1.015)	0.35	1.001 (0.998–1.004)	0.49

Note: Cox regression model was adjusted by age. Significant *p* values are shown in bold.

Abbreviations: CI, confidence interval; EDA, evidence of disease activity after IFNβ treatment; EDSS, Expanded Disability Status Scale; GFAP, glial fibrillary acidic protein; HR, hazard ratio; IFNβ, interferon-beta; NfL, neurofilament light chain; SPMS, secondary progressive multiple sclerosis.

Radiological phenotypes, but not serum biomarker levels, are associated with long-term prognosis

As a next step, an investigation of whether radiological phenotypes and biomarker levels were associated with long-term prognostic outcomes such as the time to second-line therapies, time to SPMS and time to EDSS 6.0 was carried out. For radiological phenotypes, the high inflammation phenotype was associated with a shorter time to second-line therapies compared to the low inflammation phenotype ($p = 0.04$; Figure 4a). As shown in Figure 4b, a trend towards a significantly higher proportion of patients on second-line therapies at the last follow-up visit was observed in the high inflammation phenotype (48% vs. 26% for the high and low inflammation phenotypes, respectively; $p = 0.06$). Regarding the neurodegeneration

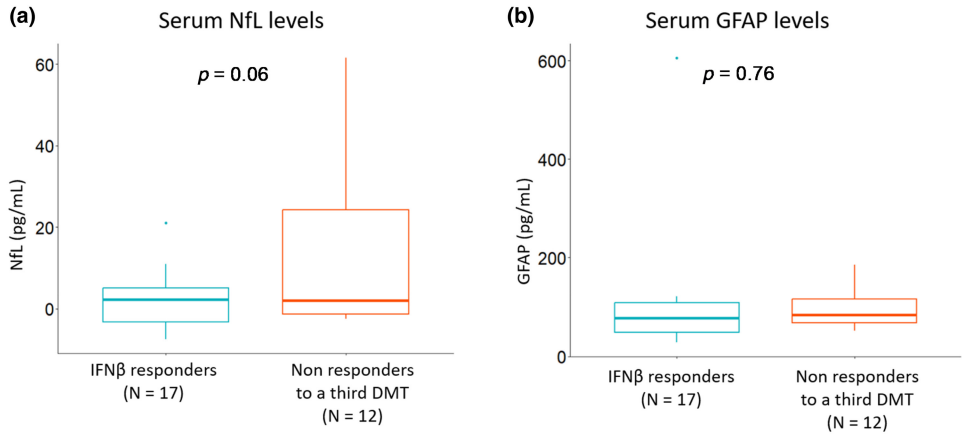

FIGURE 3 Comparison of serum biomarker levels between two extreme groups of therapeutic outcome. Box plots showing the distribution of serum NfL levels (a) and GFAP levels (b) in MS patient responders to IFNβ versus non-responders to a third disease-modifying treatment (DMT).

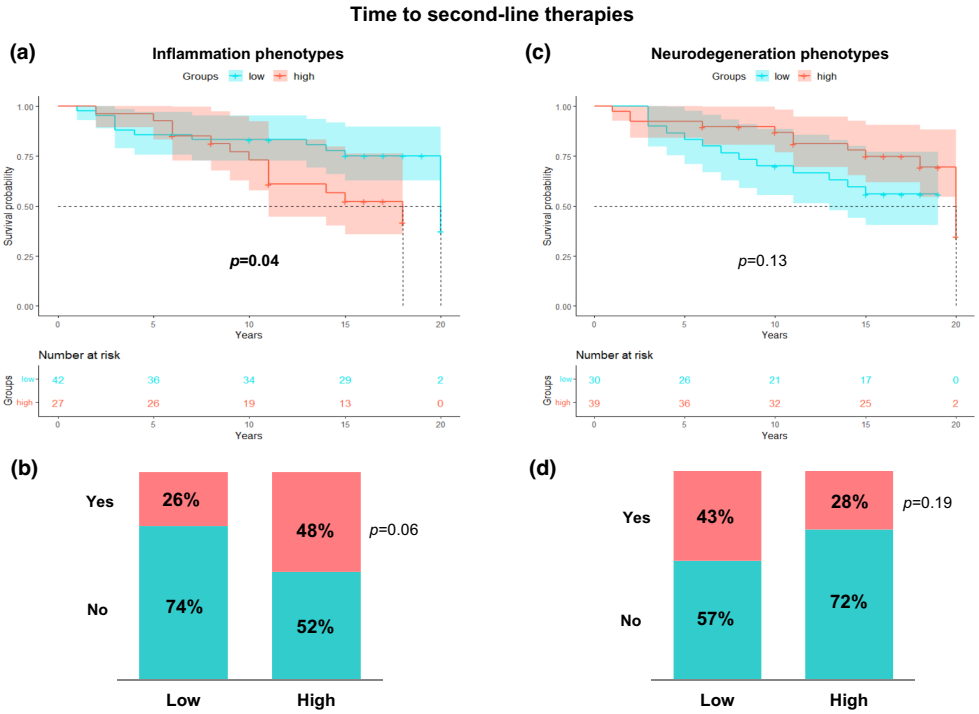

FIGURE 4 Radiological phenotypes and time to second-line therapies. Kaplan–Meier curves showing the survival of MS patients for the event initiation of a second-line therapy for high and low inflammation phenotypes (a) and high and low neurodegeneration phenotypes (c). The blue and red lines correspond to survival probability for the low and high inflammation phenotypes, respectively. Shaded areas correspond to the 95% confidence interval for each curve, and overlap between confidence intervals is represented in gray. Discontinued lines indicate median times to the event for each group. Proportion of patients from high and low inflammation phenotypes (b) and high and low neurodegeneration phenotypes (d) on second-line therapies at the last follow-up visit. "No" and "Yes" indicate patients not receiving second-line therapies and treated with second-line therapies, respectively.

phenotypes, no significant differences were observed in the time to second-line therapies or in the proportion of patients on second-line therapies at the time of last visit between patients belonging to the high and low neurodegeneration phenotypes (Figure 4c,d).

When the time to develop an SPMS disease course was evaluated, the high neurodegeneration phenotype was associated with a shorter time to SPMS compared to the low neurodegeneration phenotype ($p=0.006$; Figure 5c). In addition, the proportion of patients with SPMS at the end of the study was significantly higher in the high neurodegeneration group (46% vs. 19% for the high and low neurodegeneration phenotypes, respectively; $p=0.01$) (Figure 5d). In contrast, the time to SPMS and the proportion of patients with SPMS at the end of the study were comparable between patients from high and low inflammation phenotypes (Figures 4b and 5a).

For the outcome time to EDSS 6.0, a trend towards a shorter time to reach an EDSS of 6.0 was observed in patients belonging to the high neurodegeneration phenotype compared to patients from the low neurodegeneration phenotype ($p=0.09$; Figure 6c), whereas the time to EDSS 6.0 did not differ between patients from high and low inflammation phenotypes (Figure 6a). Likewise, the proportion of patients with EDSS 6.0 at the end of follow-up was comparable between patients with high and low inflammation and neurodegeneration phenotypes (Figure 6b,d).

Regarding biomarkers, high baseline serum NfL or GFAP levels were not associated with an increased risk for time to second-line therapies, time to develop an SPMS disease course or time to reach an EDSS of 6.0 (Table 2).

DISCUSSION

In clinical practice, the difficulty of predicting future disease trajectories and early optimization of treatment decisions remain a major challenge [1]. This situation has become even more complex in recent years due to the substantial increase of the therapeutic arsenal to treat MS patients [13, 14]. Currently, the characteristics of the baseline MRI (number, topography and activity of the lesions) constitute the main prognostic factors of disease activity [15]. However, the extreme variability of individual disease courses [1] and the lack of radiological neurodegeneration parameters applicable in clinical practice [16] make it difficult to predict long-term disease outcomes. In this context, serum biomarker quantification is likely to provide additional information relevant to understanding the pathophysiology of MS and better identifying patients at increased risk of disease severity and long-term disability [17, 18]. In a previous study conducted by our group, a cohort of RRMS patients was classified into different MRI phenotypes characterized by high and low inflammation and high and low neurodegeneration in order to correlate radiological phenotypes with specific blood transcriptomic patterns [11]. Here, the aim was first to assess whether MRI phenotypes correlated with serum levels of NfL and GFAP. In blood samples collected in proximity to the brain MRI scans used for classification of patients according to radiological phenotypes, serum NfL levels were associated with the high inflammation phenotype, whereas serum GFAP levels correlated with the high neurodegeneration

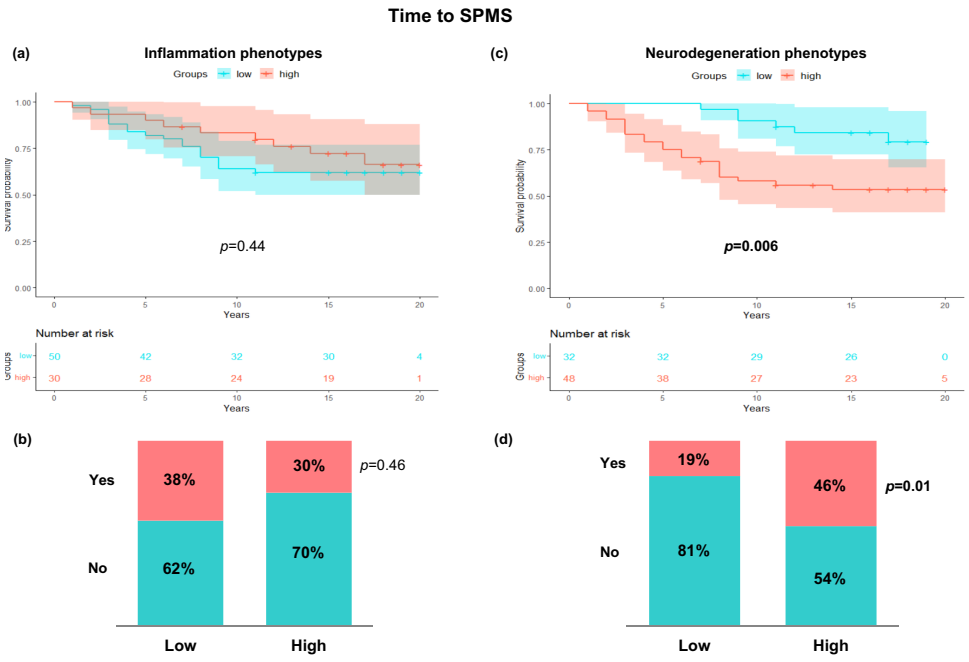

FIGURE 5 Radiological phenotypes and time to SPMS. Kaplan–Meier curves showing the survival of MS patients for the event development of SPMS for high and low inflammation phenotypes (a) and high and low neurodegeneration phenotypes (c). The blue and red lines correspond to survival probability for the low and high inflammation phenotypes, respectively. Shaded areas correspond to the 95% confidence interval for each curve, and overlap between confidence intervals is represented in gray. Proportion of patients from high and low inflammation phenotypes (b) and high and low neurodegeneration phenotypes (d) with SPMS at the end of the study. "No" and "Yes" indicate patients not developing SPMS and patients with SPMS, respectively.

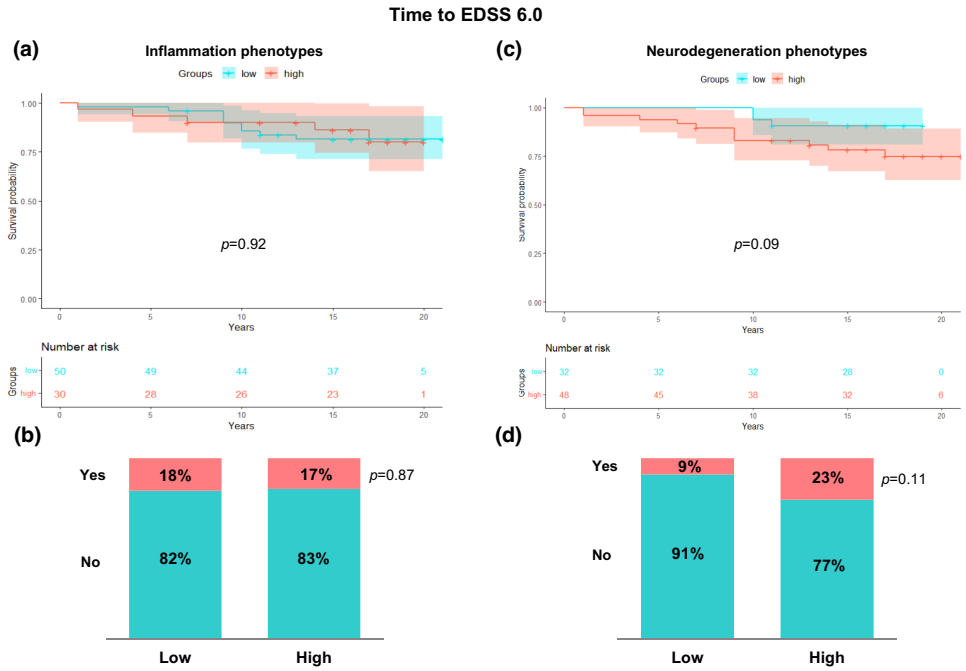

FIGURE 6 Radiological phenotypes and time to EDSS 6.0. Kaplan–Meier curves showing the survival of MS patients for the event reaching EDSS 6.0 for high and low inflammation phenotypes (a) and high and low neurodegeneration phenotypes (c). The blue and red lines correspond to survival probability for the low and high inflammation phenotypes, respectively. Shaded areas correspond to the 95% confidence interval for each curve, and overlap between confidence intervals is represented in gray. Proportion of patients from high and low inflammation phenotypes (b) and high and low neurodegeneration phenotypes (d) with EDSS 6.0 at the end of the study. "No" and "Yes" indicate patients who do not reach a EDSS 6.0 and patients who do, respectively.

phenotype. These are in agreement with previous studies reporting predominant associations of serum NfL levels with inflammation MS outcomes [19–21], and of serum GFAP levels with disability progression disease outcomes [22, 23].

In a second part of the study, the aim was to evaluate the prognostic role of inflammation and neurodegeneration radiological phenotypes and serum biomarker levels on treatment response outcomes, as well as long-term disease outcomes after a mean follow-up of 17 years. Considering that our study cohort corresponded to the baseline of IFNβ treatment, and all patients included in the study were treated with IFNβ, the response to this drug was evaluated first. The high inflammation phenotype was associated with a suboptimal response to IFNβ evaluated by the time to EDA. In concordance with these findings, high serum NfL levels, which correlated in our study with the high inflammation phenotype, were also more likely to be associated with disease recurrence in a shorter period of time after treatment onset with IFNβ. Previous studies have also reported associations between serum NfL levels and the therapeutic response to DMT [23–25]. Taking advantage of the long follow-up of our study cohort, the response to treatment in two extreme groups of treatment responders (IFNβ responders vs. non-responders to a third DMT) was also explored, and it was observed that higher NfL levels at the start of IFNβ treatment were associated with a trend towards worse response and therefore lower probability of remaining free of disease activity.

Perhaps a closer clinical-radiological monitoring would be advisable in those patients who start moderately effective therapies with high baseline serum NfL levels and/or high inflammatory activity in the baseline MRI, given the greater risk of reactivation of the disease in the short term and therefore a suboptimal response to treatment.

When evaluating long-term outcomes, only radiological phenotypes seemed to play a prognostic role in the disease. As has been reported before [15], it was found that the high inflammation phenotype was associated with a shorter time to use of second-line therapies. In addition, a greater proportion of patients belonging to this radiological phenotype required a second-line treatment at the end of the study. Maybe this is related to a common process of therapeutic switch based mainly on inflammatory rather than neurodegenerative MRI markers, the latter being very difficult to evaluate in clinical practice.

For long-term disability outcomes, associations were restricted to the high neurodegeneration phenotype, and patients belonging to this group were characterized by a shorter time to develop an SPMS disease course. Likewise, a higher proportion of patients from the high neurodegeneration phenotype had the SPMS clinical form at the time of the last follow-up visit. For the outcome time to EDSS 6.0, although a trend for shorter time was observed in the high neurodegeneration phenotype, the association did not reach statistical significance most probably due to the low frequency of patients achieving this outcome. In contrast to MRI phenotypes, serum biomarker levels were not associated with long-term prognostic factors after a mean follow-up of 17 years. Although a number of studies have shown that blood and cerebrospinal fluid NfL levels may predict future worsening of disability in MS patients, few studies analyzed long-term disability outcomes

such as conversion to SPMS and time to EDSS 6.0. In a study by Manouchehrinia et al. [26], blood NfL levels were not consistently associated with the risk of reaching a sustained EDSS of 6.0 or with the risk of conversion to SPMS after a median follow-up of 5 years. In another study by Uphaus et al. [27], blood NfL levels were associated with conversion to SPMS after a median follow-up of 6 years. Regarding GFAP, one study reported an association between cerebrospinal fluid GFAP levels and time to reach an EDSS of 6.0 in a univariable but not multivariable analysis after a mean follow-up of 12 years [28].

Our study has as limitation the relatively small sample size. Another limitation relates to the radiological measures used for classification of inflammation that includes only Gd-enhancing lesions and not new/enlarged T2 lesions, which is also considered a marker of inflammatory activity. As an additional limitation of the study, patients were included who started IFNβ, a currently little-prescribed treatment. In this context, our results should be confirmed in patients treated with the new disease-modifying drugs. However, as a strength, the selection of a homogeneous prospective cohort of patients with RRMS who started IFNβ as first treatment in all cases, with extensive follow-up, allowed us to assess short- and long-term disease outcomes. Another strength of the study is the definition of MRI phenotypes according to radiological sequences available in clinical practice.

In conclusion, radiological phenotypes defined by various degrees of inflammation and neurodegeneration have prognostic implications in treatment response and long-term disease outcomes and correlate with serum levels of NfL and GFAP respectively. Baseline serum NfL levels were associated with the treatment response outcomes but not with long-term prognostic outcomes such as time to second-line therapies, SPMS and EDSS 6.0.

AUTHOR CONTRIBUTIONS

Luciana Midaglia: Writing—original draft; methodology; writing—review and editing; data curation; conceptualization; formal analysis. Alex Rovira: Methodology; writing—review and editing; conceptualization; resources. Berta Miró: Formal analysis; writing—review and editing. Jordi Río: Writing—review and editing. Nicolás Fissolo: Methodology; writing—review and editing. Joaquín Castilló: Methodology; writing—review and editing. Alex Sánchez: Formal analysis; writing—review and editing. Xavier Montalban: Writing—review and editing. Manuel Comabella: Conceptualization; writing—review and editing; methodology; formal analysis; supervision; data curation; investigation; resources.

CONFLICT OF INTEREST STATEMENT

Luciana Midaglia has nothing to disclose. Alex Rovira serves on scientific advisory boards for Novartis, Sanofi-Genzyme, Synthetic MRI, Roche, Biogen, TensorMedical, Bristol Myers and OLEA Medical, and has received speaker honoraria from Sanofi-Genzyme, Merck-Serono, Bayer, Teva Pharmaceutical Industries Ltd, Novartis, Roche and Biogen. Berta Miró has nothing to disclose. Jordi Río has received speaking honoraria and personal compensation for participating on advisory boards from Biogen-Idec, Genzyme, Merck-Serono, Mylan, Novartis, Roche, Teca and Sanofi-Aventis. Nicolás Fissolo has nothing to disclose. Joaquín Castilló has nothing to disclose. Alex Sánchez has nothing to disclose. Xavier Montalban has received speaking honoraria and travel expenses for participation in scientific meetings, has been a steering committee member of clinical trials or participated in advisory boards of clinical trials in the past with Actelion, Amirall, Bayer, Biogen, Celgene, Genzyme, Hoffmann-La Roche, Novartis, Oryzon Genomics, Sanofi-Genzyme and Teva Pharmaceutical. Manuel Comabella has received compensation for consulting services and speaking honoraria from Bayer Schering Pharma, Merk Serono, Biogen-Idec, Teva Pharmaceuticals, Sanofi-Aventis and Novartis.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Luciana Midaglia https://orcid.org/0000-0001-7981-2228
Alex Rovira https://orcid.org/0000-0002-2132-6750
Jordi Río https://orcid.org/0000-0003-4546-7627
Manuel Comabella https://orcid.org/0000-0002-2373-6657

REFERENCES

1. Oh J, Vidal-Jordana A, Montalban X. Multiple sclerosis: clinical aspects. *Curr Opin Neurol*. 2018;31:752-759.
2. Reich DS, Lucchinetti CF, Calabresi PA. Multiple sclerosis. *N Engl J Med*. 2018;378:169-180.
3. Gafson A, Craner MJ, Matthews PM. Personalised medicine for multiple sclerosis care. *Mult Scler*. 2017;23:362-369.
4. Dobson R, Giovannoni G. Multiple sclerosis—a review. *Eur J Neurol*. 2019;26:27-40.
5. Midaglia L, Sastre-Garriga J, Montalban X. Clinical monitoring of multiple sclerosis patients by means of digital technology, a field in the midst of a revolution. *Rev Neurol*. 2021;73:210-218.
6. Wattjes MP, Ciccarelli O, Reich DS, et al. 2021 MAGNIMS-CMSC-NAIMS consensus recommendations on the use of MRI in patients with multiple sclerosis. *Lancet Neurol*. 2021;20:653-670.
7. Håkansson I, Tisell A, Cassel P, et al. Neurofilament light chain in cerebrospinal fluid and prediction of disease activity in clinically isolated syndrome and relapsing–remitting multiple sclerosis. *Eur J Neurol*. 2017;24:703-712.
8. Ziemssen T, Bhan V, Chataway J, et al. Secondary progressive multiple sclerosis: a review of clinical characteristics, definition, prognostic tools, and disease-modifying therapies. *Neurol Neuroimmunol Neuroinflamm*. 2022;10:e200064.
9. Gaetani L, Salvadori N, Lisetti V, et al. Cerebrospinal fluid neurofilament light chain tracks cognitive impairment in multiple sclerosis. *J Neurol*. 2019;266:2157-2163.
10. Ehrenberg AJ, Khatun A, Coomans E, et al. Relevance of biomarkers across different neurodegenerative diseases. *Alzheimers Res Ther*. 2020;12:56.
11. Comabella M, Cantó E, Nurtdinov R, et al. MRI phenotypes with high neurodegeneration are associated with peripheral blood B-cell changes. *Hum Mol Genet*. 2016;25:308-316.
12. Bielekova B, Kadom N, Fisher E, et al. MRI as a marker for disease heterogeneity in multiple sclerosis. *Neurology*. 2005;65:1071-1076.
13. Montalban X, Gold R, Thompson AJ, et al. ECTRIMS/EAN guideline on the pharmacological treatment of people with multiple sclerosis. *Mult Scler*. 2018;24:96-120.

14. Rae-Grant A, Day GS, Marrie RA, et al. Practice guideline recommendations summary: disease-modifying therapies for adults with multiple sclerosis: report of the guideline development, dissemination, and implementation Subcommittee of the American Academy of Neurology. *Neurology*. 2018;90:777-788.
15. Tintore M, Rovira A, Rio J, et al. Defining high, medium and low impact prognostic factors for developing multiple sclerosis. *Brain*. 2015;138:1863-1874.
16. Filippi M, Preziosa P, Barkhof F, et al. Diagnosis of progressive multiple sclerosis from the imaging perspective: a review. *JAMA Neurol*. 2021;78:351-364.
17. Domingues RB, Fernandes GBP, Leite F, Senne C. Neurofilament light chain in the assessment of patients with multiple sclerosis. *Arq Neuropsiquiatr*. 2019;77:436-441.
18. Benkert P, Meier S, Schaedelin S, et al. Serum neurofilament light chain for individual prognostication of disease activity in people with multiple sclerosis: a retrospective modelling and validation study. *Lancet Neurol*. 2022;21:246-257.
19. Khalil M, Teunissen CE, Otto M, et al. Neurofilaments as biomarkers in neurological disorders. *Nat Rev Neurol*. 2018;14:577-589.
20. Kapoor R, Smith KE, Allegretta M, et al. Serum neurofilament light as a biomarker in progressive multiple sclerosis. *Neurology*. 2020;95:436-444.
21. Barro C, Healy BC, Liu Y, et al. Serum GFAP and NfL levels differentiate subsequent progression and disease activity in patients with progressive multiple sclerosis. *Neurol Neuroimmunol Neuroinflamm*. 2022;10:e200052.
22. Abdelhak A, Foschi M, Abu-Rumeileh S, et al. Blood GFAP as an emerging biomarker in brain and spinal cord disorders. *Nat Rev Neurol*. 2022;18:158-172.
23. Kuhle J, Kropshofer H, Haering DA, et al. Blood neurofilament light chain as a biomarker of MS disease activity and treatment response. *Neurology*. 2019;92:e1007-e1015.
24. Akgün K, Kretschmann N, Haase R, et al. Profiling individual clinical responses by high-frequency serum neurofilament assessment in MS. *Neurol Neuroimmunol Neuroinflamm*. 2019;6:e555.
25. Sejbaek T, Nielsen HH, Penner N, et al. Dimethyl fumarate decreases neurofilament light chain in CSF and blood of treatment naïve relapsing MS patients. *J Neurol Neurosurg Psychiatry*. 2019;90:1324-1330.
26. Manouchehrinia A, Stridh P, Khademi M, et al. Plasma neurofilament light levels are associated with risk of disability in multiple sclerosis. *Neurology*. 2020;94:e2457-e2467.
27. Uphaus T, Steffen F, Muthuraman M, et al. NfL predicts relapse-free progression in a longitudinal multiple sclerosis cohort study. *EBioMedicine*. 2021;72:103590.
28. Martínez MA, Olsson B, Bau L, et al. Glial and neuronal markers in cerebrospinal fluid predict progression in multiple sclerosis. *Mult Scler*. 2015;21:550-561.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Midaglia L, Rovira A, Miró B, et al. Association of magnetic resonance imaging phenotypes and serum biomarker levels with treatment response and long-term disease outcomes in multiple sclerosis patients. *Eur J Neurol*. 2024;31:e16077. doi:10.1111/ene.16077